

Sprint 2 – Video Upload & AI Processing

Sprint Number: 2

Sprint Status: READY FOR EXECUTION

Sprint Duration: 10–14 Days

Methodology: Scrum

Sprint Goal:

Implement the core video processing functionality, including video uploads, caption extraction, translation, subtitle generation, text-to-speech (TTS) conversion, audio merging, and download. By the end of this sprint, approved users should be able to upload videos and receive processed outputs with AI-generated audio and translated subtitles.

User Stories Included:

- US-1: Video Upload API: Upload videos with validations and temporary storage
- US-2: Caption Extraction: Extract captions or fallback to speech-to-text
- US-3: Caption Translation: Translate captions to requested language
- US-4: Subtitle Generation: Generate .srt or .vtt subtitle files with timestamps
- US-5: Text-to-Speech (TTS): Generate audio for translated captions
- US-6: Audio-Video Merging: Merge TTS audio with original video
- US-7: User Video History: Return last 3 processed videos with TTL enforcement

Technical Scope (Sprint 2 Only):

Included:

- Video upload handling
- Caption extraction (subtitles or speech-to-text)
- Translation API integration
- Subtitle file generation (.srt/.vtt)
- Text-to-speech (TTS) audio generation
- Audio-video merging using FFmpeg
- Temporary storage & TTL enforcement
- Updated database schema for videos

Explicitly Excluded:

- Public deployment scaling
- Long video processing (>10s)
- AI enhancements beyond TTS & translation

Database Design (Sprint 2):

Videos Collection Fields: - userId (reference) - originalVideoUrl - processedVideoUrl - captionsJson - subtitlesFileUrl - audioFileUrl - language - status (uploaded/processing/complete) - expiresAt - createdAt
Indexes: - userId - TTL index on expiresAt - status for quick querying

API Design (Sprint 2):

POST /video/upload - Upload user video (verifyJWT + isApproved)
POST /video/:id/process - Process video (captions -> TTS -> merge)
GET /video/:id/download - Download processed video (verifyJWT)
GET /video/history - Get last 3 processed videos (verifyJWT)

Data Lifecycle Rules:

- Every processed video has an expiresAt timestamp
- MongoDB TTL index automatically deletes expired records
- Temporary assets in storage are deleted after 3 hours
- Users are warned to download videos within 3 hours

Definition of Done (DoD):

- Users can upload videos
- Captions are extracted and translated
- Subtitles are generated and synced with video
- TTS audio is merged with video
- Processed video is downloadable
- Video metadata stored with TTL enforcement

Known Limitations (Accepted):

- Limited video duration (5–10 seconds)
- Free-tier cloud storage only
- Manual download required within TTL
- Only approved users can process videos

Sprint Outcome:

After Sprint 2, the system will have:
- Video upload and AI processing pipeline - Translated captions and TTS audio merged with video - Secure temporary storage with TTL - Updated database schema for video tracking

Sprint 2 Conclusion:

Sprint 2 delivers the AI-powered video processing functionality, building on the secure foundation from Sprint 1. The system will be ready for multi-language processing, subtitle management, and TTS features.